Czech speech rhythm and the rhythm class hypothesis

Dancovicova, Jana ; Dellwo, Volker

Abstract: While a number of languages have been classified as either syllable- or stress-timed, the case of Czech remains unclear. In this paper we make predictions about Czech rhythm on the basis of our analysis of syllable complexity in recorded samples of Czech. The results on syllable complexity show mixed features. This is reflected in the classification of Czech rhythm using rhythm measures based on durational variability of consonantal and vocalic intervals.

Czech Speech Rhythm and the Rhythm Class Hypothesis

Jana Dankovičová & Volker Dellwo

University College London
j.dankovicova@ucl.ac.uk, v.dellwo@ucl.ac.uk

ABSTRACT

While a number of languages have been classified as either syllable- or stress-timed, the case of Czech remains unclear. In this paper we make predictions about Czech rhythm on the basis of our analysis of syllable complexity in recorded samples of Czech. The results on syllable complexity show mixed features. This is reflected in the classification of Czech rhythm using rhythm measures based on durational variability of consonantal and vocalic intervals.

Keywords: Prosody, speech rhythm, Czech.

1. INTRODUCTION

It has been demonstrated that on the acoustic level rhythmic classes (stress-, syllable-, or mora-timing) are represented in the variability of consonantal (c) and vocalic (v) intervals ([7], [9]). These acoustic measures compare either the durational variability between consecutive c- and v-intervals (Pairwise Variability Index, henceforth: PVI; [7]), the overall durational variability (standard deviation of c- and v-intervals, henceforth: ΔC and ΔV respectively; [9]) or simply the percentage over which speech is vocalic (henceforth: %V; all measures are further discussed under 4). It has further been demonstrated for some of these measures that they are perceptually salient ([9]).

Rhythmic theories based on the variability of cv-interval durations follow the rationale that the syllable structure of languages is responsible for their rhythmic characteristics ([1], [2], [7], [9]). In this respect languages traditionally classified as stress-timed show a more complex syllable structure than syllable-timed languages. This is usually characterized by the permission of complex c-clusters on the consonantal level and vowel reductions on the vocalic level.

In the present experiment we attempt to test this rationale by analyzing Czech, a language which has been traditionally classified as syllable-timed [8] but which, as shown in [6], can be classified also as stress-timed, depending on the measure used. First we analyze the syllable complexity of Czech and compare it to languages that have been shown to belong to stress- and syllable-timed classes by [7] and [9]. On the basis of this comparison we then make predictions as to how the rhythmic measures based on c- and v-interval variability should behave. Finally, we will test these predictions on real speech.

2. EXPERIMENTAL METHOD

The current experiment makes use of the BonnTempo corpus, a speech database created for the analysis of rhythm in a number of languages ([5]). In this database speakers have been recorded reading a text of about 7 clauses in their native language (about 90 syllables in each language), translated from German into all languages under investigation by educated native speakers. For the current experiment this text was translated into Czech.

BonnTempo contains samples of read speech at different intended speech rates (i.e. speakers were instructed to read at a normal rate, slowly and fast). For the present study, however, only the material with intended normal speech rate is considered.

Nine native Czech speakers from Prague (5 female, 4 male; age 23-70 yrs) were recorded reading the Czech passage. From the BonnTempo corpus speakers representing stress-timing (German, n=15 and English, n=7) and speakers representing syllable-timing (French, n=5, and Italian, n=3) were chosen for comparison.

All speech material in BonnTempo was manually labelled according to syllable and cv-interval durations ([5]). The Czech data was labelled in the same way by the first author of this paper.

3. CZECH SYLLABLE COMPLEXITY

In this section we compare the syllable complexity in the Czech reading passage with the syllable complexity of the passages in the stress- (English, German) and syllable-timed (French, Italian) languages.

The results are presented in Figure 1. The pie-charts show that the text excerpts for the languages traditionally classified as syllable-timed (French and Italian) have a much lower percentage of complex syllables than those classified as stress-timed (English and German). In French about 80% and in Italian about 90% of syllables consist of a single vowel or a consonant and vowel (with a clear preference for CV rather than VC). Only 10% of the

Italian syllables are CVC or CCV, while in the French text there is a slightly higher percentage of CVC syllables (18.3%) and a negligible number of CVCC and CCVC syllables (1.2% and 2.3%, respectively). In English and German the syllable complexity is considerably higher, with German showing the most complex syllables. In German only 35% of the syllables are V, CV or VC; in English it is slightly more than 50%. However, both languages contain a considerably higher number of complex syllables than syllable-timed languages.

The Czech data reveal that this language lies in between the two classes of languages. Simple syllable structures (V, CV, VC) are represented by 65% of syllables, which is higher than stress-timed English and German, but clearly lower than syllable-timed French and Italian.

If measures of variability in c-interval durations are mainly influenced by syllable complexity, then we can hypothesize that Czech has somewhat lower c-interval variability than English and German but higher variability than French and Italian.

With regard to variability in v-interval durations, it has been argued that the main factor contributing to the division between stress- and syllable-timed languages is vowel reduction, since stress-timed languages allow it and syllable-timed languages do not allow it to the same extent ([2], [7] and [9]). Consequently, it has been claimed, stress-timed languages show a higher variability in v-interval durations ([7], [9]).

In the case of Czech the situation is again different. Czech does not allow vowel reduction, however, it has a phonological distinction between short and long vowels, with the durational ratio between them being about 1:2 [8]. In the passage used for the present paper 16% of vowels were long and 84% short (this fits in well with the published data on Czech in [11]). Thus from the point of view of vowel reduction, Czech should be syllable-timed and measures of v-interval variability should reflect this by placing Czech close to French and Italian. However, the phonological long/short distinction may well introduce significant variability. Since the influence of systematic use of phonological length on unit durations in such a context has not yet been studied, to our knowledge, it is difficult to make precise predictions about what v-variability measures should show.

4. CZECH RHYTHM ANALYSIS

Rhythm has been analyzed according to two different models: (a) Ramus et al. [9] (henceforth: Ramus model) and (b) Grabe & Low [7] (henceforth: Grabe & Low model).

4.1. Ramus model

Ramus proposes that c- and v-interval variability are best captured by the measures %V and ΔC. %V represents the overall percentage over which speech is vocalic. Languages allowing v-reduction (most stress-timed languages) are assumed to have a higher variability in v-interval durations as a consequence and thus a smaller overall percentage over which speech is vocalic (i.e. smaller %V, see [9] for the underlying rationale). ΔC measures variability in c-interval durations in terms of standard deviations. This measure is supposed to be directly affected by syllable complexity. It is argued ([7], [9]) that languages traditionally classified as stress-timed allow complex c-clusters and thus show a higher durational variability in c-interval durations which in return results in a higher ΔC.

[1] demonstrated that the standard deviations of c-intervals (ΔC) are strongly negatively correlated with speech rate (higher rate leads to shorter mean c-intervals, resulting in smaller standard deviations). [9] normalized for speech rate by carefully selecting sentences of roughly equal duration and equal number of syllables. For data not controlled in this way [4] suggested using the variation coefficient of c-interval durations (henceforth: varcoC; (ΔC*100)/mean c-interval). Since sentence durations and number of syllables in the data at intended normal speech rate in BonnTempo and in our Czech version vary highly across and within speakers, we used varcoC rather than ΔC in order to measure c-interval variability.

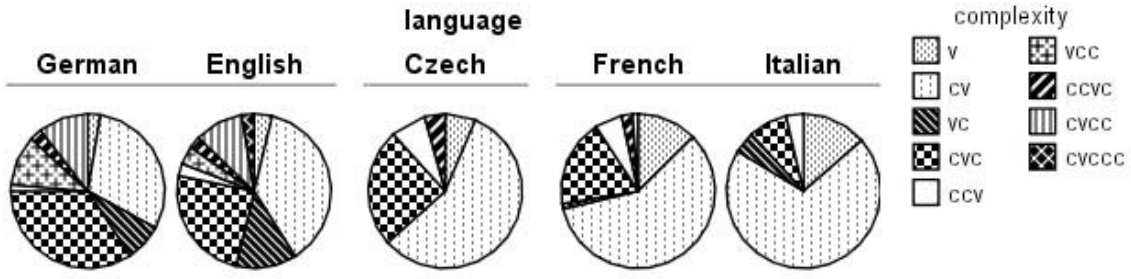
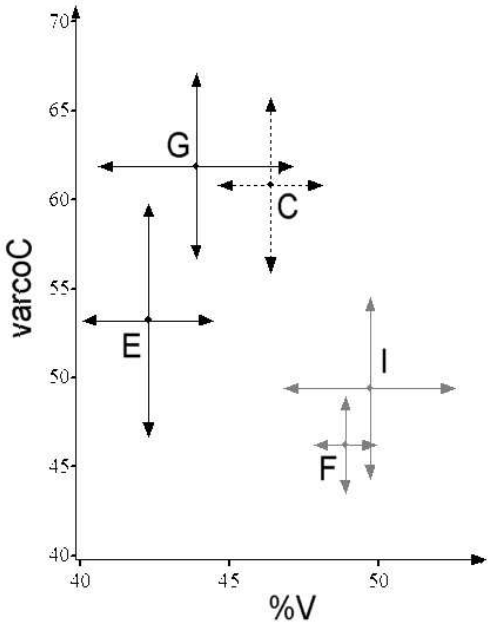


Figure 1: Syllable complexity for Czech, English, French, German, and Italian

Figure 2 shows the results for varcoC and %V plotted against each other (the mean values for each language were averaged across all speakers; arrows represent standard deviations).

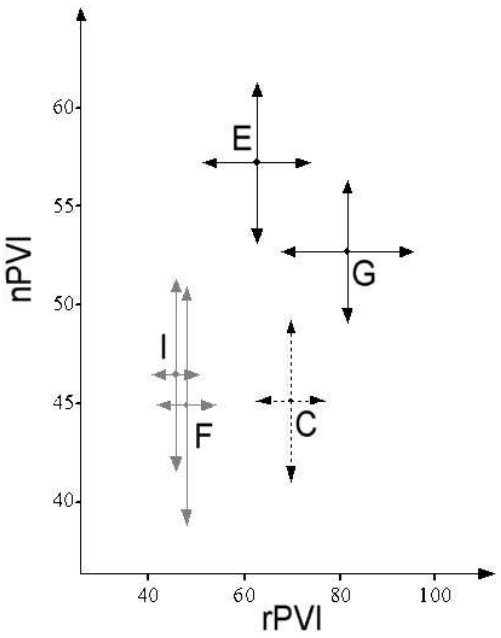
On the %V axis the graph shows that traditional syllable-timed languages (French, Italian) are well separated from stress-timed languages (English, German). An ANOVA with language as a factor and %V as dependent variable yields a highly significant effect ($F[4,40]=9.24$, $p<0.001$). A post-hoc test (Tukey) shows that languages within the same rhythm class (English/German vs. French/Italian) are not significantly different from each other ($p>0.05$). All between rhythm class comparisons are significant at least at the $p<.05$ level. The post-hoc test for Czech shows that this language differs significantly only from English ($p=.023$) but not from any other language.

Figure 2: Mean values of varcoC plotted against %V for Czech (C), English (E), German (G), French (F) and Italian (I). Arrows represent standard deviations**.**



C-interval variability in our data does not distinguish the traditional stress-timed/syllable-timed languages as clearly as %V. The only significant distinction was found between German and all other languages. All other comparisons yielded non-significant results. However, when languages are grouped into rhythm classes (as done by [9]), a significant difference between stress- (English, German) and syllable-timed (French, Italian) groups is obtained. With respect to Czech the test reveals that this language is significantly different from all languages but German.

Figure 3: nPVI mean values plotted against rPVI mean values for Czech (C), English (E), German (G), French (F) and Italian (I). Arrows represent standard deviations.



4.2. Grabe & Low

Grabe & Low [7] argue that the phonotactic /phonological complexity of c- and v-intervals is reflected in durational differences between consecutive intervals. These differences are registered by the so called Pairwise Variability Index (PVI) which calculates an average of the sum of consecutive interval differences. Grabe & Low distinguish between a raw PVI (rPVI) applied to c-interval variability and a speech rate normalized PVI (nPVI) for v-intervals which divides each v-interval difference by the sum of the respective two intervals and is thus supposed to control for speech rate. It has been pointed out, however, that this is a very local normalization that does not take account of changes in speech rate on a level higher than two consecutive v-intervals (see [1]). More effective speech rate normalization methods for nPVI are not yet available but are currently being developed by the second author of the present paper and we expect to report on them soon.

Figure 3 shows values for nPVI plotted against rPVI (the mean values for each language averaged

across all speakers). The results are broadly in line with those obtained by Ramus model (above). On the vocalic level (nPVI) significant differences were found for all between rhythm class language comparisons while within class comparisons were non-significant. On the consonantal level (rPVI) the results are the same as for varcoC: German is significantly different from all other languages while no other significant differences can be found.

Regarding Czech, however, the data draws a somewhat different picture. According to nPVI Czech overlaps fully with Italian and French (no significant difference), while highly significant differences ($p<.005$) were found between Czech and English as well as Czech and German. In short, according to nPVI, Czech is clearly syllable-timed. The rPVI results, however, suggest a different conclusion. Here the differences between Czech and stress-timed English and German are non-significant while the differences between Czech and syllable-timed French and Italian are both significant. Thus, according to rPVI Czech is stress-timed.

5. DISCUSSION

The results presented here call into question the validity of recent measures for rhythm classification suggested in [7] and [9]. The analysis of syllable complexity suggests that Czech should be between stress- and syllable-timing on a c-interval level, as its syllable structure is overall simpler than in traditional stress-timed languages, but more complex than in syllable-timed languages. Neither rPVI nor varcoC reflect this. In fact, both measures show that Czech is closer to English and German (traditionally stress-timed) than to French and Italian (traditionally syllable-timed). Moreover, we found that, according to varcoC, stress-timed English is significantly different from German – the language in the same rhythm class – but not from either of the syllable-timed languages. One explanation may be that the link between syllable complexity and variability in c-interval durations is not as straightforward as assumed. However, another explanation may lie in the fact that some types of consonants (nasals) contribute more to c-interval variability than others, as found by [10]. We are therefore planning to investigate Czech consonantal structure in more detail and compare the use of different consonant classes across the languages under investigation.

As regards v-interval variability the Ramus and Grabe & Low measures clearly disagree. %V shows that Czech is in between syllable- and stress-timing; nPVI shows that it is clearly syllable-timed. %V best reflects our hypothesis based on the fact that Czech does not have vocalic reductions but does have phonological vowel length distinction (see above) and should therefore be in between the stress- and syllable-timed extremes. Given that %V fits our predictions better, we suggest two possible explanations for why nPVI does not reflect the prediction well: (a) since nPVI is a 'pairwise' variability measure that only captures differences between consecutive vocalic intervals, it may not be able to capture the phonological length differences; (b) nPVI has been demonstrated to be speech rate dependent ([1]) and our reading material shows inter- and intra-speaker variations of rate. It is therefore possible that the current nPVI measures are an artefact of rate. We are currently investigating a normalization procedure that uses durations relative to the total sentence duration in the hope of being able to explain some of the variability in our own data as well as in other studies ([6], [7]).

6. REFERENCES

[1] Barry, W.J., Andreeva, B., Russo, M., Dimitrova, S., Kostadinova, T. 2003. Do rhythm measures tell us anything about language type? *Proc. 15th ICPhS* Barcelona, 2693–2696.

[2] Dasher, R., Bolinger, D. 1982 On pre-accentual lengthening. *Journal of the International Phonetic Association* 12, 58-69.

[3] Dauer, R.M. 1983. Stress-timing and syllable-timing reanalysed. *Journal of Phonetics* 11, 51-69.

[4] Dellwo, V. 2006 *Rhythm and Speech Rate: A Variation Coefficient for deltaC.* Karnowski, P. & Szigeti, I. (eds.) *Language and Language-processing.* Frankfurt am Main: Peter Lang: 231-241.

[5] Dellwo, V., Steiner, I., Aschenberner, B., Dankovičová, J., Wagner, P. 2004. The BonnTempo-Corpus and BonnTempo-Tools. A database for the combined study of speech rhythm and rate. *Proc. 8th ICSLP,* Jeju Island, Korea.

[6] Duběda, T. 2004. K izosylabičnosti a izochronnosti v češtině. In: Sborník z Konference česko-slovenské pobočky ISPhS, Univerzita Karlova, Praha, Filozofická fakulta, 19-28.

[7] Grabe, E., Low, E.L. 2004. Durational variability in speech and the rhythm class hypothesis. In: Gussenhoven, C., Warner, N. (eds) *Papers in Laboratory Phonology 7*, Berlin, New York: Mouton de Gruyter.

[8] Palková, Z. 1994. *Fonetika a fonologie češtiny*. Prague: Karolinum.

[9] Ramus, F., Nespor, M., Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.

[10] Steiner, I. 2005. On the analysis of speech rhythm through acoustic parameters. In: Fisseni, B., Schmitz, H-C., Schröder, B., Wagner, P. (eds) *Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen*. Frankfurt (Main): Peter Lang, 647-658.

[11] Těšitelová, M. et al. 1987. *O češtině v číslech.* Prague: Academia.